2013

Traffic Analytics using Probabilistic Graphical Models Enhanced with Knowledge Bases

Pramod Anantharam
Wright State University - Main Campus, anantharam.2@wright.edu

Krishnaprasad Thirunarayan
Wright State University - Main Campus, t.k.prasad@wright.edu

Amit P. Sheth
Wright State University - Main Campus, amit@sc.edu

Repository Citation

Anantharam, P., Thirunarayan, K., & Sheth, A. P. (2013). Traffic Analytics using Probabilistic Graphical Models Enhanced with Knowledge Bases. *Proceedings of the 2nd International Workshop on Analytics for Cyber-Physical Systems (ACS-2013)*, 13-20.
https://corescholar.libraries.wright.edu/knoesis/143

Traffic Analytics using Probabilistic Graphical Models Enhanced with Knowledge Bases

Pramod Anantharam Krishnaprasad Thirunarayan Amit Sheth

{pramod, tkprasad, amit}@knoesis.org
Kno.e.sis - Ohio Center of Excellence in Knowledge-enabled Computing
Wright State University, Dayton, USA

Abstract

Graphical models have been successfully used to deal with uncertainty, incompleteness, and dynamism within many domains. These models built from data often ignore preexisting declarative knowledge about the domain in the form of ontologies and Linked Open Data (LOD) that is increasingly available on the web. In this paper, we present an approach to leverage such "top-down" domain knowledge to enhance "bottom-up" building of graphical models. Specifically, we propose three operations on the graphical model structure to enrich it with nodes, edges, and edge directions. We illustrate the enrichment process using traffic data from 511.org and declarative knowledge from ConceptNet. The resulting enriched graphical model can potentially lead to better predictions of traffic delays.

1 Introduction

Cyber-Physical Systems (CPS), which have tightly coupled computation, communication, and control components, are increasingly being used in various domains [27]. The observations and interactions in a CPS are characterized by: (1) *incompleteness* due to partial observation from the real world, (2) *uncertainty* due to inherent randomness involved in the sensing process (noise in case of machine sensors and bias in case of citizen sensors), and (3) *dynamism* from the ever changing and non-deterministic conditions of the physical world. Graphical models can be used to deal with incompleteness, uncertainty, and dynamism in many diverse domains such as traffic management, healthcare, system health monitoring, speech processing, image processing, and computational biology [23, 28, 1, 29, 2]. These models are built bottom-up using sensor data in most cases, i.e., the structure (conditional independence between variables) and parameters (Conditional Probability Distribution – CPD for continuous variables, or Conditional Probability Tables – CPT for discrete variables) are learned from sensor data [14]. Some of these models also allow domain experts to express their declarative knowledge of the domain in the form of value constraints and inequality constraints in the learning process [19].

Extracting structure of a graphical model from observations of a CPS is very challenging due to data sparsity, incompleteness, and difficulty in detecting causal links. However, declarative domain knowledge can obviate the need to learn everything from data. In addition, correlations derivable from data can be further consolidated if the declarative knowledge base provides evidence that it is causal in nature. For example, the influence of a music event on traffic flow may be known a priori. However, this declarative knowledge is insufficient to answer quantitative questions which are addressed by graphical model parameters.

Declarative knowledge (including causal relationships) is increasingly being published using open data standards on the Semantic Web [3]. These include knowledge bases such as ConceptNet5 [18] and many domain ontologies and data sets published on the Linked Open Data (LOD) [4] cloud. We hypothesize that leveraging such knowledge will increase the fidelity of graphical models. More importantly, it will complement structure learning algorithms of graphical models by utilizing declarative domain knowledge. Specifically, we focus on how knowledge gaps (incompleteness) can be reduced.

The contributions of this work include:

- definition of external events and internal observations that co-exist across CPS,
- extraction of graphical model structure by correlating external events (e.g., music concert) to internal observations (e.g. delays),
- definition of three operators (addition of nodes, edges, and edge directions) for complementing structure learning of graphical models using declarative domain knowledge (e.g., ConceptNet), and
- evaluation of the approach on a real world traffic

data set (511.org).

The enriched model is more encompassing of the domain variables and relationships between them, which can potentially lead to better prediction of delays.

2 Motivation

City authorities have realized the need for traffic management because current unmanaged traffic results in 1.9 billion gallons of wasted gas per year, costing drivers over $100 billion in wasted gas and time anually in U.S. alone [6]. There are large deployments of sensor networks monitoring traffic flows in many cities such as San Francisco[1], Dayton[2], New York, and Chicago[3]. 511.org is one such service that provides traffic and transit data feeds for San Francisco nine-county Bay Area. Traffic data feed, transit data feed, ride match service, driving times API, and real-time transit departures API deliver real-time data to subscribers. Traffic data feed consists of information on links (road segments and their locations), link status (current speed, volume, and travel time), and event information (active and scheduled). Examples of active events include accidents, vehicle breakdowns, and signal fault. Scheduled events include sporting events, planned road work, and music events.

For understanding traffic flow, we need to explore correlations between traffic flow variations and the events in a city. We propose a systematic approach for uncovering associations between internal observations (e.g. delays) and external events (e.g. accident) at different locations in a city. Figure 1 contains link status and event information from 511.org service along with relevant declarative knowledge from ConeptNet. The graphical structure extraction techniques [14] can extract the link between scheduled event and slow traffic given historical observations. The *directionality of the link* can be fixed using declarative knowledge from ConceptNet as shown. E.g. the causal link between *base ball game* and *traffic jam* decides the direction of the link. We use a comprehensive categorization of active and scheduled events provided by 511.org to infer the types of each event observation, e.g., *base ball game* is a scheduled event. *Data sparsity* may result in missing links in the structure extraction process, e.g., association between *traffic jam* and *time of day* may not be readily inferred. Figure 1 shows knowledge from ConceptNet that can be used to overcome the limitation of conventional structure extraction. *Incompleteness* is pervasive in all real world data, e.g., given the absence of weather data, its effect on traffic jam is not captured by structure extraction. Declarative knowledge from ConceptNet provides this information complementing structure extraction process as shown in Figure 1. In this research, we address only the incompleteness aspect out of the three challenges that exist in any CPS.

[1]http://www.511.org/
[2]http://www.buckeyetraffic.org/
[3]http://www.traffic.com/

3 Related Works

We organize related works into two categories based on their focus. The first one corresponds to general approach to build graphical models from ontologies and the second one corresponds to specific approach to traffic analytics.

3.1 Graphical Models from Ontologies Use of domain knowledge in the construction of Bayesian networks has been widely studied. Utilizing inequality and range constraints on the statistical variables in a domain is explored in [17]. These constraints, provided by domain experts, are incorporated in the learning process using a constrained EM learning algorithm which minimizes the violation of constraints. A systematic process of transforming an ontology into Bayesian network is presented for modeling oesophagus cancer [8]. Use of ontologies in construction of Object Oriented Bayesian Network (OOBN) [15] is studied by several works [16, 13, 7, 12]. These approaches leverage object oriented techniques for representation and reasoning in modeling large and complex domains.

In contrast, CPS modelling require causal knowledge in addition to subsumption and named relationships specified in the ontologies. Some of the named relationships in the ontologies may be causal in nature. We use ConceptNet which explicitly specifies causal knowledge, to enrich the structure of graphical models extracted from observational data. The principles of enrichment can extend across domains, but, for concreteness in using real world data, we chose the domain of traffic which has open real world data.

3.2 Traffic Analytics The research on traffic flow analysis can be categorized into two broad categories. The first category deals with traffic flow analysis considering only the sensor observations monitoring the road network [9, 25]. They are called internal observations since they are limited to traffic flow patterns and are agnostic to events outside the sensor network (i.e., external events). The second category is inclusive of events external to the sensor network monitoring the traffic flow. External events include active events (e.g., accidents, vehicle breakdowns) and scheduled events (e.g., sporting events, music events) [22, 10]. External events may be obtained by city authorities from sources such as

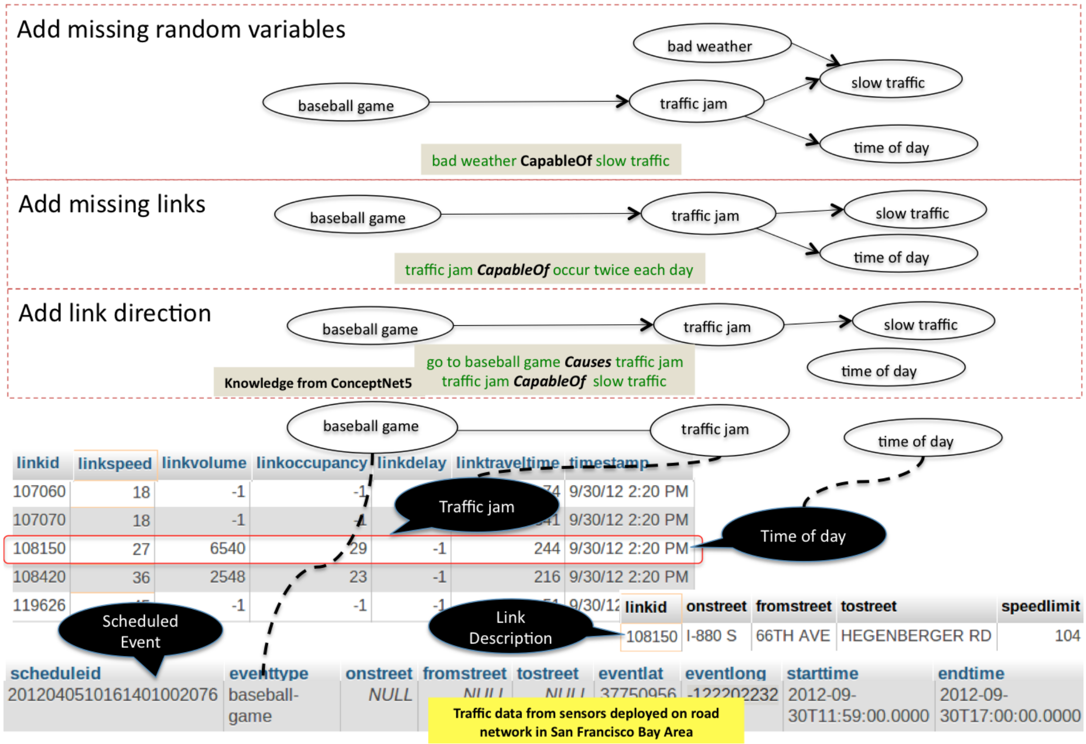

Figure 1: Extraction of correlation between slow traffic and scheduled event from real world data (511.org) and enriching graphical model with declarative knowledge from ConceptNet

511.org or by citizens in the city through social streams such as tweets.

In complement to the existing approaches, we envision broadening traffic analysis to use domain knowledge in understanding observations to derive deeper insights into the processes that govern these complex systems. The traffic flow in a city is not entirely independent of the events in a city and behavior of its inhabitants. Many of the events in the city may influence traffic. These events include inclement weather, music events, sporting events, accidents, and breakdowns. Domain knowledge and statistical approaches provide complementary benefits leading to a system that understands patterns in data (bottom-up) while leveraging existing background knowledge (top-down).

Specifically, we focus on use of declarative domain knowledge of traffic for complementing structure of graphical models. Some of the questions we try to answer regarding declarative knowledge include: Will it help us in finding missing domain variables? Will it help us in finding missing edges/relations between the variables? Will it help us in assigning missing directions of edges/relations?

4 Preliminaries

We define notations and terminologies for representing sensor observations and define abstractions over these observations which are used in rest of the paper. Recall that we categorize the traffic related observations into two major categories: (1) Internal observations, and (2) External events. Internal observations consists of observations from on-road sensors monitoring flow of traffic. They are called "internal" since these observations are internal to the road network, e.g., speed, volume, and travel time. External events constitutes all the events external to the road network that may or may not influence traffic. Scheduled and active events constitutes external events.

4.1 Notations A link is a fundamental building block of a road network. Multiple links connected back to back constitutes a road. Each link is monitored by sensors measuring speed of vehicles, volume of vehicles, and time to travel the link. Let $\langle O_s, O_v, O_t \rangle$ be the corresponding variables representing internal observations of a link. $\langle E_a, E_s, D_{week}, T_{day}, O_{delay} \rangle$ are binary variables representing active and scheduled events, day of the week, time of the day, and delay in link travel time, which are all external events. Each of these are repre-

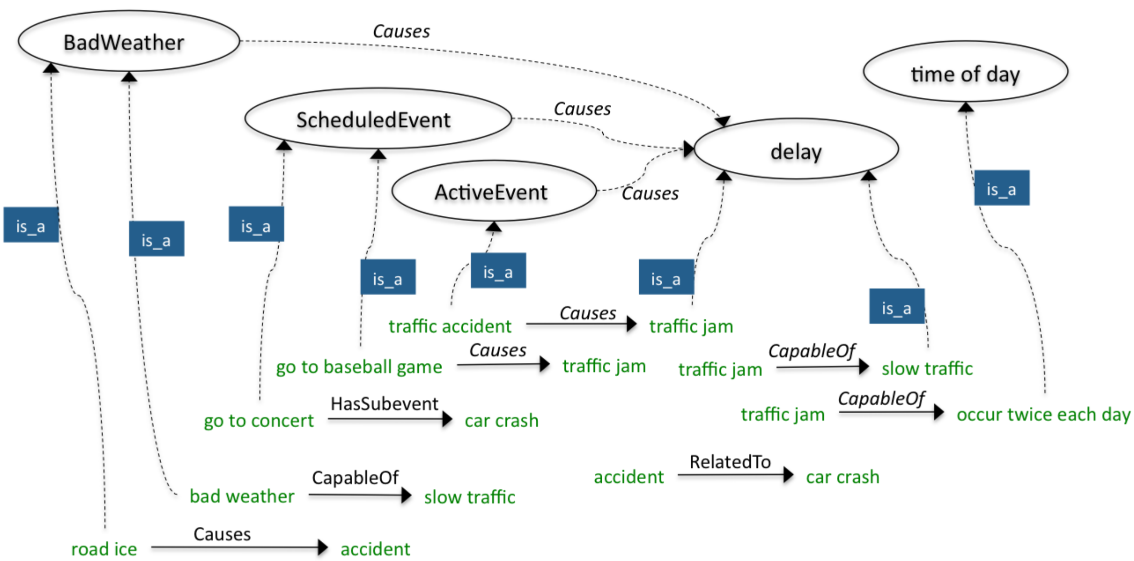

Figure 2: Domain knowledge of traffic in the form of concepts and relationships (mostly causal) from ConceptNet

sented as a random variable, whose value is unknown till we observe it. External events have start time (t_s) and the duration it persists (t_d, provided by 511.org), using which we compute the event end time (t_e). We generate a training set by processing each observation record (reported almost every two seconds) where O_s, O_v, and $O_t \in \mathbb{R}$ are sensor observations. E_a, E_s, D_{week}, T_{day}, and $O_{delay} \in \{0,1\}$, and are set to 1 depending on whether the external event is active. We compute the average travel time $T_{avg,l,h}$ for each link (l) for every hour of the day (h ranges for 24 hours) for total number of days (n days) considered for analysis. We represent the dataset created using tuples of the form $\langle O_s, O_v, O_t, E_a, E_s, D_{week}, T_{day}, O_{delay} \rangle_{timestamp}$

Domain knowledge from ConceptNet, shown in Figure 2, is represented by a set of concepts and relationships between them ($G_{declarative}$).

4.2 Problem Build a graphical model G inherent in the domain of traffic consisting of random variables $\langle O_s, O_v, O_t, E_a, E_s, D_{week}, T_{day}, O_{delay} \rangle_{timestamp}$. The structure consists of connections between random variables that specifies conditional independence statements. The parameters are the CPT or CPD learned from data. Our focus is on complementing structure extraction with declarative knowledge. Parameter learning (CPT or CPD with probability values/distributions) is out of scope of this work.

Problem 1 Extract the structure G from traffic data.

Problem 2 Use $G_{declarative}$ to modify G. This step may add (1) new random variables to G, (2) new links, and (3) link directionality.

5 Approach

We now present the details of our approach of complementing graphical model structure extraction with declarative knowledge from knowledge sources. We focus on structure extraction and propose parameter learning from declarative knowledge as a future work. Starting from the overall system architecture, this section provides details of the approach toward utilizing declarative knowledge in the construction of graphical models for CPS.

5.1 System Architecture The overall system architecture is shown in Figure 3. The raw observations from traffic data feed of 511.org are preprocessed to generate training data. The training data consists of random variables used to describe the domain of traffic which includes O_s, O_v, O_t, E_a, E_s, D_{week}, T_{day}, and O_{delay}. The Java Messaging Service (JMS) is used as a mode of subscription to 511. The XML messages are parsed to extract scheduled and active events, and link status (speed, volume, and travel time) information. These observations are further processed to generate boolean abstractions such as time of day, peak hour, and day of week. This data set is used as a training input to the structure learning algorithm which extracts dependencies between the random variables.

5.2 Preprocessing of Traffic Observations Traffic observations are processed using the Algorithm 1 to

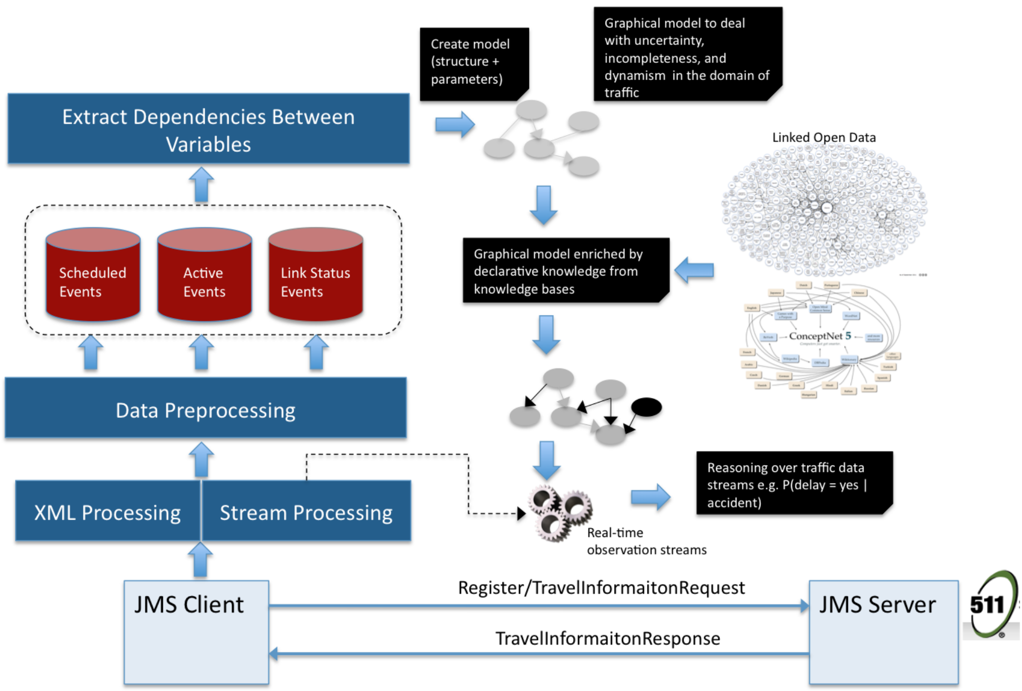

Figure 3: Traffic observation stream processing pipeline along with our approach to complement graphical models with knowledge from existing knowledge bases on the web

generate the training data, which is then used by the structure learning algorithms to extract inherent structure in the domain of traffic. There are additional fields added in the preprocessing step derived from raw sensor observations and active and scheduled event status. The raw observations are mapped to binary values to form the training data. The timestamp of the observation is used to check for presence of active events, scheduled events, day of week, time of day, and current link delay status. The corresponding boolean variable is set to 1 depending on the presence of the event at the observation timestamp.

Algorithm 1 Algorithm for preprocessing traffic data stream

for each $\langle O_s, O_v, O_t \rangle_{timestamp} \in$ TrafficStream **do**
  $E_a = 0$, $E_s = 0$, $D_{week} = 0$, $T_{day} = 0$, $O_{delay} = 0$
  if timestamp $\in$ timestamp of E_a **then**
    $E_a = 1$
  else if timestamp $\in$ timestamp of E_s **then**
    $E_s = 1$
  else if timestamp $\in$ saturday, sunday **then**
    $D_{week} = 1$
  else if timestamp $\in$ peak-hour; 7 AM to 9 AM $\vee$ 4 PM to 6 PM **then**
    $T_{day} = 1$
  else if $O_t > \mathrm{T}_{avg,l,h}$ **then**
    $O_{delay} = 1$
  end if
  $\langle O_s, O_v, O_t, E_a, E_s, D_{week}, T_{day}, O_{delay} \rangle_{timestamp}$
end for

5.3 Structure Learning for Graphical Model

Figure 4 shows the structure extracted using observations $\langle O_s, O_v, O_t, E_a, E_s, D_{week}, T_{day}, O_{delay} \rangle_{timestamp}$. There are undirected links which need to be converted to directed links before the parameters are learned. We explored many structure extraction approaches including constraint-based, score-based, and hybrid algorithms. Constraint-based learning algorithms verify the conditional independence between variables with ideas derived from [24]. Score-based learning algorithms evaluate the structure and compute the goodness-of-fit score. Hybrid algorithms combine the aspects of constraint-based and score-based algorithms for structure extraction. These algorithms are implemented as part of the *bnlearn* [26] package of R [11], which is a statistical data analysis tool. Within constraint based, we used Grow-Shrink [20], Incremental Association, Fast Incremental Association [30], and Interleaved Incremental Association [32]. In score-based we explored Hill-Climbing and Tabu Search. Max-Min Hill-Climbing and Restricted Maximization [31] were used to extract structure as part of hybrid learning approach. There are other constraint-based approaches such as Chow-Liu [5] and ARACNE [21] that learn the correlation graph (without any direction) between random variables. Chow-Liu approach gave us the best structure (i.e. intuitively most satisfactory) that described the traffic domain and we chose this for further refinement using declarative knowledge.

5.4 Knowledge Base to Complement Structure Learning The structure learned from Chow-Liu's approach resulted in a correlation graph shown in upper part of Figure 4. While correlations are entirely data-driven and may or may not support semantic interpretation, the causal links which are more valuable, are much harder to extract. We hypothesize that, we can use declarative knowledge from knowledge base such as ConceptNet to infer causal links in the structure extracted by structure learning algorithms. The graph G in general is composed of nodes (N), edges (E), and direction (D) of edges i.e., G = ⟨ N, E, D ⟩. We distinguish the extracted graph from the declarative knowledge graph $G_{declarative} = \langle N_{dk}, E_{dk}, D_{dk} \rangle$

We focus on the causal and subsumption knowledge related to traffic from ConceptNet. We propose three operations on the graph structure G (result of structure extraction that encapsulates the sorts of updates that can be performed to graphical models as a result of using declarative knowledge) using the knowledge base, $G_{declarative}$: (1) Add missing random variables $G_n \leftarrow \phi_{addnode}(G, G_{declarative})$, (2) Add missing links $G_e \leftarrow \phi_{addedge}(G_n, G_{declarative})$, and (3) Set directionality of the links $G_{complete} \leftarrow \phi_{adddirection}(G_e, G_{declarative})$. We assume that the subset of relevant declarative knowledge is given to us. The first step is to unify the variables represented in G and $G_{declarative}$ using subsumption and then perform the three operations.

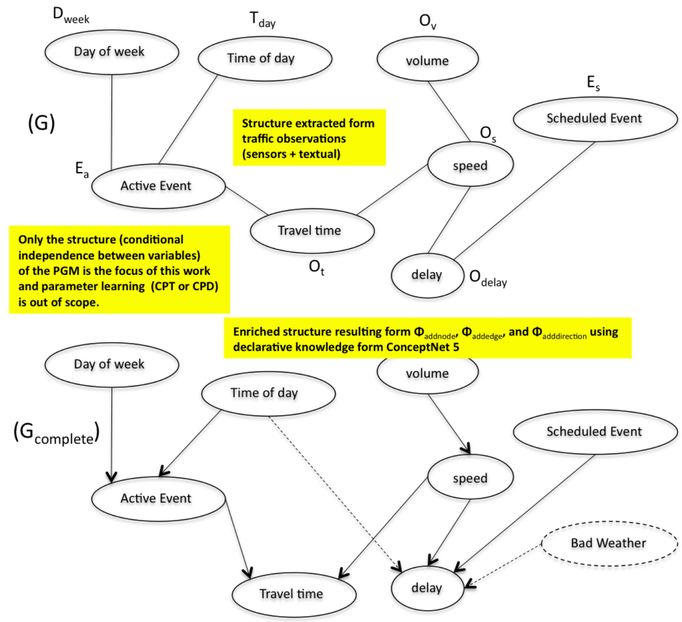

Figure 4: Top part of the figure depicts the structure extracted from traffic observations and the bottom part has the enriched structure (using declarative knowledge)

5.4.1 Node Addition $\phi_{addnode}$: Due to lack of instrumentation, there may be missing or incomplete observations from a domain. Structure learning algorithms cannot account for such missing observations. We leverage domain knowledge to add missing random variables to G. From $G_{declarative}$, we add all the nodes missing in G, i.e., N = N ∪ (N_{dk} - N).

5.4.2 Edge Addition $\phi_{addedge}$: Due to data sparsity, the structure extracted may have missing links between different random variables. Purely data driven approaches may have this limitation. We use domain knowledge to add missing links to G. From $G_{declarative}$ we add all the edges missing in G, i.e., E = E ∪ (E_{dk} - E).

5.4.3 Edge Directionality $\phi_{adddirection}$: The structure learning algorithms use information theoretic and statistical techniques to extract correlations between the random variables. While correlations give us links between the random variables, they do not provide the directionality of the links. A domain expert looking at the links can decide the directionality of the links to form causal links. We exploit such a declarative domain knowledge from knowledge bases to form causal links. For all the links in G without a direction (missing entry in D), we look for corresponding directionality in D_{dk} and add it to D. $G_{complete}$ is a graph that is a directed acyclic graph which is obtained by leveraging domain knowledge.

6 Evaluation

We describe the dataset used for evaluation and exemplify the enrichment of the graphical model using declarative knowledge from knowledge bases.

6.1 Dataset Description We evaluate the approach over traffic sensor observations collected for a week from October 14th to 20th from San Francisco Bay Area. We accumulated over three million sensor observations during this period along with 20 active events and 7 scheduled events as shown in Table 1. We chose this interval since there were many external events and some missing internal observations. Such a condition usually exists in analyzing real world data.

First, we extract the structure from data and next we use the operations $\phi_{addnode}$, $\phi_{addedge}$, and

Table 1: Dataset used for evaluation of structure extraction

Day	SensorObs	ActiveEvents	ScheduledEvents
10/14/12	109258	0	5
10/15/12	547305	14	1
10/16/12	1024256	0	0
10/17/12	495386	0	0
10/18/12	0	0	1
10/19/12	928374	0	0
10/20/12	0	0	0

$\phi_{adddirection}$ to enrich it. The enriched structure is compared with the initial structure extracted from data using qualitative comparison.

The upper part of Figure 4 shows the structure (correlations) that exist between different variables in the traffic domain. The bottom part of Figure 4 has the enriched structure resulting from the three operations we proposed.

Qualitative Comparison: The structure extracted from data gives correlations between the variables in the domain of traffic. The directionality of the edges are unknown (upper part of Figure 4). The correlations seems intuitive, such as time of day and day of week are related to active events (e.g. peak hours and week days may have more accidents and breakdowns), active events affect travel time (e.g. an accident may increase the travel time of a link), and scheduled events affect delays (e.g. music event at a location may cause traffic delays).

The bottom part of Figure 4 has directions assigned to all the links. Day of the week and time of the day may influence the occurrence of active events. Scheduled events and bad weather may cause delays. The dotted node is the missing variable in the structure extracted from traffic data. The dotted edge between time of day and delay was a missing edge added to the network. Figure 5 validates this relationship obtained from declarative knowledge. The enriched model can account for the effect of weather (e.g. poor visibility) on delays, which is indeed realistic in real world situations and may lead to accurate prediction of delays.

7 Conclusions

Cyber physical systems (e.g., road traffic network) are characterized by observations spanning textual (e.g., incident report) and numerical observations (e.g., speed of vehicles). The internal observations in the domain of traffic such as speed, volume, and travel time are affected by external traffic related events. We proposed a novel approach for leveraging domain knowledge (e.g., by reusing causal knowledge from ConceptNet) in extraction of dependencies between variables in

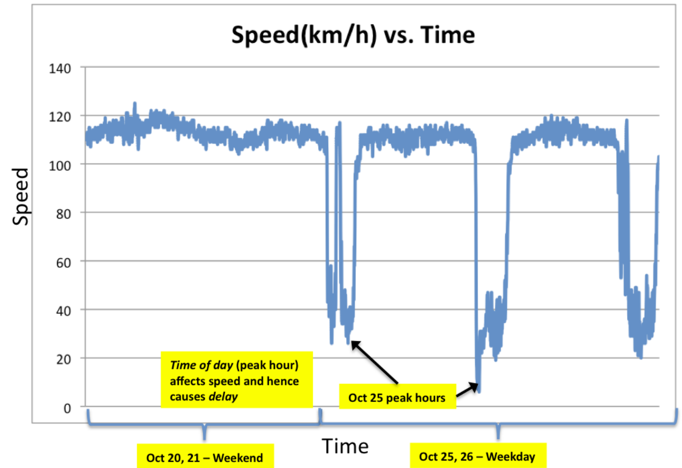

Figure 5: Variations of speed over a week (Oct 20th – 26th, 2012) for a link with linkid = 101170 – validating the correlation between *time of day* and *delay* in travel time

the domain of traffic. Three operators (that add missing nodes, add missing edges, and assign edge directions) for enrichment of graphical models using declarative knowledge were defined. We exemplified the enrichment process using real world traffic data from 511.org and concepts from ConceptNet. Our evaluation showed that combining graphical models with qualitative information present in declarative knowledge provides much richer domain model for reasoning. Specifically, it allows us to combine two complementary sources of information: (i) quantitative and correlation-based knowledge automatically synthesized bottom-up from traffic data, and (ii) manually curated qualitative and causal knowledge available top-down. The declarative knowledge from ConceptNet and the structure of the graphical model are both qualitative and at a comparable level of abstraction, relative to the quantitative information extracted via the parameters. So, we have restricted ourselves to extracting only the structure in the form of conditional dependencies from data and ignore parameter estimation in the form of conditional probability values.

8 Acknowledgments

We would like to thank Dr. Payam Barnaghi for insights on knowledge extraction from CPS.

References

[1] E.M. Airoldi, *Getting started in probabilistic graphical models*, PLoS Computational Biology, 3 (2007), p. e252.

[2] J. Bilmes and G. Zweig, *The graphical models toolkit: An open source software system for speech and time-series processing*, in Acoustics, Speech, and Signal Processing (ICASSP), 2002 IEEE International Conference on, vol. 4, IEEE, 2002, pp. IV–3916– IV–3919.

[3] C. Bizer, T. Heath, and T. Berners-Lee, *Linked data-the story so far*, International Journal on Semantic Web and Information Systems (IJSWIS), 5 (2009), pp. 1–22.

[4] C. Bizer, T. Heath, K. Idehen, and T. Berners-Lee, *Linked data on the web (ldow2008)*, in Proceeding of the 17th international conference on World Wide Web, ACM, 2008, pp. 1265–1266.

[5] C. Chow and C. Liu, *Approximating discrete probability distributions with dependence trees*, Information Theory, IEEE Transactions on, 14 (1968), pp. 462–467.

[6] DOT, *A new economic analysis of infrastructure investment, a report prepared by the department of the treassury, with the council of economic advisers*, 2012.

[7] S. Fenz, A.M. Tjoa, and M. Hudec, *Ontology-based generation of bayesian networks*, in Complex, Intelligent and Software Intensive Systems, 2009. CISIS'09. International Conference on, IEEE, 2009, pp. 712–717.

[8] E.M. Helsper and L.C. van der Gaag, *Building bayesian networks through ontologies*, in ECAI, vol. 2002, 2002, p. 15th.

[9] A. Hofleitner, R. Herring, P. Abbeel, and A. Bayen, *Learning the dynamics of arterial traffic from probe data using a dynamic bayesian network*, ITS, IEEE Transactions on, 13 (2012), pp. 1679 –1693.

[10] E.J. Horvitz, J. Apacible, R. Sarin, and L. Liao, *Prediction, expectation, and surprise: Methods, designs, and study of a deployed traffic forecasting service*, arXiv preprint arXiv:1207.1352, (2012).

[11] R. Ihaka and R. Gentleman, *R: A language for data analysis and graphics*, Journal of computational and graphical statistics, 5 (1996), pp. 299–314.

[12] M.B. Ishak, P. Leray, N.B. Amor, et al., *Ontology-based generation of object oriented bayesian networks*, in Proceedings of the 8th Bayesian Modelling Applications Workshop, 2011, pp. 9–17.

[13] Mouna Ben Ishak, Philippe Leray, and Nahla Ben Amor, *A two-way approach for probabilistic graphical models structure learning and ontology enrichment*, in KEOD'11, 2011, pp. 189–194.

[14] D. Koller and N. Friedman, *Probabilistic graphical models: principles and techniques*, MIT press, 2009.

[15] D. Koller and A. Pfeffer, *Object-oriented bayesian networks*, in Proceedings of the Thirteenth Conference on Uncertainty in Artificial Intelligence, Morgan Kaufmann Publishers Inc., 1997, pp. 302–313.

[16] H. Langseth and T.D. Nielsen, *Fusion of domain knowledge with data for structural learning in object oriented domains*, The Journal of Machine Learning Research, 4 (2003), pp. 339–368.

[17] W. Liao and Q. Ji, *Learning bayesian network parameters under incomplete data with domain knowledge*, Pattern Recognition, 42 (2009), pp. 3046–3056.

[18] H. Liu and P. Singh, *Conceptneta practical commonsense reasoning tool-kit*, BT technology journal, 22 (2004), pp. 211–226.

[19] D. Madigan, J. Gavrin, and A.E. Raftery, *Eliciting prior information to enhance the predictive performance of bayesian graphical models*, Communications in Statistics-Theory and Methods, 24 (1995), pp. 2271–2292.

[20] D. Margaritis, *Learning Bayesian network model structure from data*, PhD thesis, University of Pittsburgh, 2003.

[21] A.A. Margolin, I. Nemenman, K. Basso, C. Wiggins, G. Stolovitzky, R.D. Favera, and A. Califano, *Aracne: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context*, BMC bioinformatics, 7 (2006), p. S7.

[22] M. Miller and C. Gupta, *Mining traffic incidents to forecast impact*, in Proceedings of the ACM SIGKDD International Workshop on Urban Computing, ACM, 2012, pp. 33–40.

[23] N. Oliver and A.P. Pentland, *Graphical models for driver behavior recognition in a smartcar*, in Intelligent Vehicles Symposium, 2000. IV 2000. Proceedings of the IEEE, IEEE, 2000, pp. 7–12.

[24] Judea Pearl and T.S. Verma, *A theory of inferred causation*, 1991.

[25] C. Rotsos, J. Van Gael, A.W. Moore, and Z. Ghahramani, *Probabilistic graphical models for semi-supervised traffic classification*, in Proceedings of the 6th International Wireless Communications and Mobile Computing Conference, ACM, 2010, pp. 752–757.

[26] M. Scutari, *Learning bayesian networks with the bnlearn r package*, Journal of Statistical Software, (2009), pp. 1–22.

[27] L. Sha, S. Gopalakrishnan, X. Liu, and Q. Wang, *Cyber-physical systems: A new frontier*, Machine Learning in Cyber Trust, (2009), pp. 3–13.

[28] David J. Spiegelhalter, *Bayesian graphical modelling: a case-study in monitoring health outcomes*, Journal of the Royal Statistical Society: Series C (Applied Statistics), 47 (1998), pp. 115–133.

[29] A.N. Srivastava and J. Han, *Machine learning and knowledge discovery for engineering systems health management*, vol. 22, Chapman & Hall, 2011.

[30] I. Tsamardinos, C.F. Aliferis, A. Statnikov, and E. Statnikov, *Algorithms for large scale markov blanket discovery*, in Proceedings of the sixteenth international Florida artificial intelligence research society conference, 2003, pp. 376–381.

[31] I. Tsamardinos, L.E. Brown, and C.F. Aliferis, *The max-min hill-climbing bayesian network structure learning algorithm*, Machine learning, 65 (2006), pp. 31–78.

[32] Sandeep Yaramakala and Dimitris Margaritis, *Speculative markov blanket discovery for optimal feature selection*, in ICDM, IEEE, 2005, pp. 809–812.